Humanists and Linked Data

Report on Experiments with Pundit

Steffen Hennicke[1], Gerold Tschumpel[1], Christian Morbidoni[2], Alessio Piccoli[2], Kristin Dill[3], Stefan Gradmann[4]

(1) Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Deutschland

(2) Semedia, Università Politecnica della Marche, Piazza Roma, 22, 60121 Ancona, Italia

(3) Österreichische Nationalbibliothek, Josefsplatz 1, 1015 Wien, Österreich

(4) Literary Studies Research Unit, Leuven, Blijde-Inkomststraat 21, 3000 Leuven, België

Table of Contents

1. Report on Experiments

As part of Task 3.4, three experiments have been conducted with the semantic annotation application Pundit and additional components such as Korbo for simple vocabulary management and instance data creation and faceted browsers in the second half of 2014. While the experiments conducted in the context of the Wittgenstein Incubator focused on the usability of Pundit (cf. D1.4), the experiments reported on here were designed to provide empirical input and add a practical bottom-up perspective to the more theoretical and top-down research regarding the functional primitives and Scholarly Operations as well as the "reasoning" (cf. introduction to the Deliverable).

The research interest of the experiments was to investigate how interpretative approaches of humanists can be operationalised in the particular context of Linked Data and Pundit and its components. For this purpose, humanists were confronted in real-life working contexts with the formal and explicit approach of Linked Data and semantic annotation.

The experiments particularly focused on the Scholarly Activities *annotating* and *visualising* both of which are seen as being pivotal to most humanists research activities. The aim was to investigate how these two activities materialise in different real-life use cases (cf. Scholarly Operations) focusing on interpretative approaches of humanists which have no prior knowledge of Linked Data or semantic annotations tools such as Pundit.

The principal topical and temporal horizon for the experiments was historical sciences and Contemporary History (19th/20th century).[1] Based on this precondition, during the first half of 2014, more than 70 historians and teachers mostly at German history

[1] The main reason for choosing this topical and temporal orientation was that the two organisers of the experiments are trained historians which was expected to facilitate and support the experiments.

departments and similar institutions were contacted and asked whether they personally or in the context of a seminar with students were willing to participate in the experiment. About 30 responses were received of which about 20 were positive and interested. From there we began investigating 8 different use cases and different topical orientations. After an initial round of deliberation with potential participants we chose 3 use cases for implementation.[2]

The **three distinct use cases** chosen for the experiment belong to the historical-archival domain. The first use case which has been created in cooperation with the Berlin-Brandenburgische Akademie der Wissenschaften (BBAW) and the Fachhochschule Potsdam (FHP) focused on the editorial and archival sciences. The use case created with the Georg-Eckert-Institut (GEI) focused on educational history. The use case devised with the historical seminar of the Humboldt-Universität zu Berlin (UBER) focused on visual history and the didactics of history, i.e. using tools such as Pundit for teaching history to students.

All three experiments and their respective use cases targeted the same overall research questions and were conducted by employing the same principal methodological approach. This principle plan for the experiments had been approved by the DHAB in its 4th meeting.

The experiments investigated the following principal **research questions**:

- How can genuine research questions and interests be operationalised within the context of Linked Data and the particular context of Pundit?
    - How do non-experts deal with Linked Data concepts and approaches?

[2] The decision was based on different aspects such as the availability of appropriate digital material, the possibility to organise a seminar with students, and available time of the teachers and historians to invest into the experiments.

- How are the Scholarly Activities annotation, visualisation, and interpretative modelling reflected as Scholarly Operations?
    - Which “statements” do humanists consider necessary in their particular use case?
- Which potential do they see in Linked Data and Pundit for applications in the humanities?

The **methodological approach** and set-up of the experiments consisted of the following three principal stages:

1. Series of preparatory meetings with teachers/lecturers

   The first step in the preparation of the experiment was the identification of research questions or interests which were relevant to the given context of the use case. Then, the identified research questions and interests were operationalised for the context of Pundit and Linked Data in the form of simple annotation vocabularies. These meetings were part of the experiment and included open discussions and hand-written protocols as means of data recording.

2. Series of workshops and work at home

   The second phase of the experiments consisted of series of meetings with the students in the form of several workshops. In general, the participants were introduced to the particular seminar topics by their teachers. Afterwards, they were introduced to Linked Data and Pundit and the prepared annotation vocabularies were discussed. The students then worked with Pundit and the annotation vocabulary in the workshops as well as at home. During these workshops, data was collected by observational and interrogative means and recorded in hand-written protocols. At the end of each use case an extensive questionnaire was filled out by the participants.

3. Follow-up meetings with teachers

   The follow-up meetings were meant to reflect on the previous workshops and in particular to provide input on the future application scenarios and general advantages and disadvantages of Linked Data and Pundit. These meetings consisted of semi-structured discussions and were recorded by hand-written protocols.

In sum, data has been collected by open **interviews**, i.e. open discussions with the participants during every phase of the experiments, **observation** during the workshops including the research data created and collected in the notebooks, and via a common **questionnaire** at the end of each experiment. This questionnaire contained overall questions pertaining to the potential, shortcomings, advantages and disadvantages of Linked Data and Pundit as well as dedicated sections for each experiment.

All three use case followed the principle set-up outlined above, however, each one also focused on additional and distinct aspects. The section “Use Cases” will introduce the topic and conduct of each use case in a broader perspective. Input for this section is based on the interviews and observations. The section “Questionnaire” will then take a more comprehensive perspective on the experiments, Linked Data, and Pundit by reporting on the most important results from the questionnaire. The section “Conclusion” will summarise the most relevant results and provide recommendations for engaging humanists with Linked Data in a fruitful and productive way.

2. Use Cases

In the following, all three use cases will be presented. The particular topics will be introduced and the preparatory phase and the conduct of the experiments discussed. In

particular, we will comment on the creation and application of the annotation vocabularies.

For the experiments holds that we had to operationalise each use case in the context of the current capabilities and state of Pundit and its components. With regard to the vocabulary creation that means we refrained from creating hierarchical class and property definitions since Pundit currently does not display these hierarchies. Therefore, we created flat annotation vocabularies. Also, we did not reuse classes or properties from existing ontologies for efficiency and time reasons. With regard to visualisation we also decided due to time constraints to utilize, and at the same time test, the available generic visualisation through the in-built faceted browser in Ask as well as simple custom-build faceted browser similar to the Wittgenstein Faceted Browser.[3]

1.1 Fachhochschule Potsdam (FHP)

The first use case stems from the discipline of Editorial and Archival Sciences. The workshop was held in collaboration with Markus Schnöpf, from the Berlin-Brandenburgische Akademie der Wissenschaften, who is involved in the Digital Humanities, in particular Digital History and Editorial and Archival Sciences and is associated to a lectureship for a seminar at the Fachochschule Potsdam (FHP). All of the participants of the workshop, 15 in number, were Bachelor students attending the seminar “Editionstechniken” and the workshop was intended to complement the seminar with respect to digital techniques and methodology, in particular the application of Linked Data and the Pundit environment for Editorial and Archival Sciences. The Institut für Bibliotheks- und Informationswissenschaft of the Humboldt-Universität zu Berlin hosted the workshop on two dates, 22 August and 12 September 2014.

[3] Cf. http://metasound.dibet.univpm.it/dm2e/ajax-solr-master/examples/wab/

The problem statement was if the creation of editorial guidelines for archival material are possible in a Linked Data context and if such guidelines in the form of a simple annotation vocabulary can be successfully applied in the context of Pundit. The topical focus of the use case was the crisis of July 1914. Diplomatic primary sources from the political archive of the German foreign ministry[4] were selected in advance by the teacher.

During the preparation of the workshop, a basic editorial guideline was devised in cooperation between the teacher and the organisers of the experiments. The guideline specified to mark-up (1) all textual phenomena concerning the structure of the documents, such as title, signatures, or remarks, (2) different scripts (“Hände”), (3) basic metadata such as author, date, or provenance, (4) persons and (5) places mentioned in the text, and (6) events referred to. Additionally, if possible, the participants were asked to relate the entities to existing entities from authority files like Virtual International Authority File (VIAF),[5] Geonames,[6] and DBpedia.[7] In the case of Geonames and VIAF, this was done manually by replacing URIs of instances in Korbo.

Based on this guideline, a simple core annotation vocabulary was prepared for the workshop. During the workshop, a group of students was asked to extend the vocabulary based on the guideline. One reason was that the conceptual creation of an editorial guideline, in this case the validation and possible extension of the prepared guideline, was part of the students seminar task. The other reason was that the particular focus of this use case was on the creation of a suitable vocabulary by the students themselves: are students able to translate the conceptual framework of editorial and archival sciences to a suitable vocabulary?

[4] http://www.archiv.diplo.de/

[5] http://viaf.org/

[6] http://www.geonames.org/

[7] http://dbpedia.org/

The vocabulary group extended the core vocabulary with text phenomena found in selected primary sources and based on the input by the other participants of the workshop. Considering the short timespan (1,5 weeks) for the creation of the extension and specification of the vocabulary worked well: The most difficult conceptual problem for the participants was the differentiation between class and instance. A minor issue was the level of abstraction of the properties: several properties could have been subsumed under more general ones. Another difficulty for the participants was to determine whether they had created a comprehensive set of entities for the description of relevant phenomena in the primary sources.

After the group finished their specification, the lecturers implemented the vocabulary in Pundit. The participants then each selected one primary source and started working with the vocabulary based on the aforementioned guidelines. They were asked to document their work as well. Part of the working instructions were that the students had to not only provide a label for new instances but also a scope note describing the meaning of the new instance.

The appropriateness of the vocabulary has been proved by the fact that the other participants were able to apply the vocabulary during the workshops and that they did not ask for any significant additions during their work and also not in the questionnaire. Lastly, the actual triples created in the notebooks show that the participants did in fact successfully apply the editorial guidelines to the primary sources.

The translation of a simple editorial guideline to a RDF vocabulary proved to be possible during the preparatory phase and during the revision and extension phase during the workshop.

All the necessary or essential statements, as identified in the context of this seminar, for editorial work are factual statements, and, consequently, were easily representable in

the triple structure. Examples of such factual statements, in contrast to more interpretative statements, are statements about structural text phenomena such as pages, signatures, titles etc., and, factual statements about the contents such as the author, topic, addressee etc.

A principal conceptual issue, which is not specific to the use case and which did not pose any practical problems during the workshop, are statements about the exact provenance of a digitised text: Is the provenance of the digital text the same as the analogue one? Another principal issue is the exact semantics of statements about the phenomena in the text: should a statement about the author of a text have the complete document and/or the proper name in the text as its subject? What exactly are we talking about when we refer to phenomena in the text, which are represented in the digital copy of this text?

The translation process also showed the chance to avoid over specification (über-diplomatisch) while still retaining the potential to extend and specify the vocabulary as needed. Whether editorial scientists should mark-up more or less phenomena in texts and whether user of edited texts do profit from over detailed markup remains open. The technical requirements towards the formalisations and implementation of an annotation vocabulary in Pundit, however, constitutes an opportunity to rethink these issues especially regarding open digital and networked working environments.

1.2 Georg-Eckert-Institut (GEI)

The second use case stems from the discipline of Educational History. In this case, only one participant, a trained historian from the Georg-Eckert-Institut (GEI) in Braunschweig, took part in the experiment. Furthermore, the experiment was not organised as a series of two workshops but the annotation work with Pundit was conducted over the course of several weeks in August and September 2014.

Preparatory meetings and a follow-up meeting were held as in the case of the other two experiments.

The overall topic of the use case came from the project "World of Children"[8]. The research question was how children have been influenced and educated in their formative years in school. Investigating the formative years of adults yields important insights into how they think and write on the discourse on modernity. Textbooks are semi-official documents that were read by wider parts of the Germans during their formative years. With this material we try to find the representations of the world and the nation and the description of historical processes that were offered by the state to its future citizens. So, we search for representations of the nation and the globalised world. Also, we look for representations of change, crisis, religious conflict, social change and similar events.

The goal is to identify various topoi and their connotation and presentation in different kinds of school books: Which topoi appear in the different kinds of school books? How are they connotated and in which context are they put? These topoi will be compared over time, i.e. around 1850 and around 1900 in order to assess which and how specific topoi and their connotation change and which new ones appear or old ones disappear. For example, "nation", "globalisation", or "forming of the nation" are topoi which are discussed very differently in protestant and catholic school books. The connotation connected with topoi also differ: For example, "the Kaiserreich" is associated with backwardness, and "the Kaiserreich" prevents the founding of a German nation. Lastly, the question is if these topoi and connotations can be grouped into specific "images of others" ("Fremdbilder") and "images of oneself" ("Selbstbilder")? The study is a

[8] http://www.dipf.de/de/dipf-aktuell/pdf-aktuelles/presseinformationen/pm-2014/PM_2014_29_04ProjektstartWeltderKinder.pdf

qualitative analysis on the small scale. It may serve as a framework for subsequent and more extensive analysis (re-usability).

For the purpose of this experiment, the participant chose school books from the Digital Library of the GEI which are also available in the DM2E repository. Annotations were made on the digitised pages of the chosen books.

This experiment additionally focused specifically on the aspect of reasoning. In contrast to the other use cases, the participant was explicitly asked to use ASK[9], a faceted browser for exploring annotations in notebooks created with Pundit, to try to explore new hypotheses based on filtering annotations. The results of this part of the experiment have been reported in the section on reasoning (cf. “Reasoning with Reasoning”).

During several meetings the research question and approach was developed and a basic annotation vocabulary of properties created. The participant then worked independently with Pundit over the course of several weeks in August and September 2014.

Since this particular use case had only one participant working over a longer time period, it was possible to create a detailed description of the actual working process. The first steps formally belong to the source critique in the historical methodology and included, in this particular case, the semi-random reading of the source material. Reference points for in-depth-view of the material were subheadings. Interesting text sections were annotated and annotated fragments labelled. Then, triples were created about the source material regarding factual statements about the author, publication date, title, etc. After that, triples of second order were created which identified important historical persons and events. These subjects were then combined with either places,

[9] http://demo-cloud.ask.thepund.it/

dates or states. Lastly, references were created to material outside of the corpus such as DBpedia in order to explain parts of the material to non-experts.

The focus of the vocabulary devised for the experiment is on the properties, i. e. the expressivity of the vocabulary stems from the different types of relations between the phenomena within the school books. In contrast to the first use case the phenomena marked up in the text in this use case are subject to interpretation in an extent that is significantly different from each other. The marked up phenomena consist mainly of connotations and subtle undertones. Irrespective of the phenomena as entities themselves, this enables the correlation of relations between them in a substantially more flexible way.

Yet, these phenomena can be expressed with the simple triple structure: Since the focus of these phenomena is on the type of relation between two entities, expressive properties can be utilised to express these relations.

The properties carry specific interpretations themselves merging several distinct statements into one property. For example, the property “is positively modern connotated with” is a complex statement embodying a hypothesis about the expected semantics of a text, i.e. between two entities. The semantic of the property expresses that in the context of a particular text an entity, the subject of triple, is presented as modern in a positive way by being discussed in the context of another entity, the object, which stands for modernity in a positive sense. For example, the “Reichsgründung” (founding of the Reich) “is positively modern connotated with” “Wirtschaftseinheit” (unity of economy). Both, the object and the subject are implicitly topoi.[10] This is a way to reduce complex semantics of interpretation to a reasonable level of abstraction.

[10] In so far, the property should have as its domain and range a class ‘Topoi’.

In addition to that, interpretation is a potentially unending process of recurring succession, i.e. the certainty of particular statements made at one point during the working process may increase or decrease when confronted with new knowledge. This does not only apply to the material that is subject to interpretation but to the vocabulary as a means of interpretation itself. Thus it may be necessary to introduce new distinctions to annotate, beyond the annotation of what is immanent to the text initially worked with.

During the follow-up meeting, the participant stressed this fact again that there are layers or different levels of interpretation which may build upon each other. These levels should be clearly distinguishable in a network of statements, for example, statements immediately referring to the context or reality of the text phenomena and statements referring to high-level interpretation, transcending the immediate context of the text. A network of interpretative statements may continuously evolve, creating new hypothesis but probably subsequently also demanding new properties and classes.

1.3 Humboldt-Universität zu Berlin (UBER)

The third use case stems from the discipline of visual history and the didactics of history. The workshop was held in collaboration with Sabine Moller, from the Department of History of the Humboldt-Universität zu Berlin, who is focusing on Didactics of History. 15 students from Humboldt-Universität zu Berlin (UBER) took part in the experiment. All participants had a background in history even though they came from different institutes and participated in different programs (Bachelor and Master). The experiment was part of the seminar “Fotografie und Geschichte digital” (photography and digital history) which was held over the course of two weeks with four all-day meetings at the end of October and beginning of November 2014. The seminar included presentations by the lecturers on the topic of visual history and historical analysis of historical photographs along with introductions to Pundit and Linked Data.

The particular problem statement of the experiment was two-fold: if Pundit and Linked Data are able to support learning critical analysis of digitised historical photographs and if Linked Data can be used to enable historical critique of visual source documents. In that regard, the seminar was located at the intersection of Didactics of History and Visual History.

During the preparatory phase, we decided to refrain from having the participants create a vocabulary on their own but to prepare a ready-to-use vocabulary for the seminar. The reason was that the translation of the requirements of a historical analysis to a RDF vocabulary demanded more time and effort than would have been feasible for the students in the seminar.

During the workshop the students first learned about how to analyse historical photographs and then were asked to compare what they learned to the prepared vocabulary. The vocabulary was slightly modified and extended during the course of the first workshop day. After a general introduction to Linked Data and exercises with Pundit, the students searched and selected their own photographs from the Web in order to analyze them with Pundit by applying the annotation vocabulary. This phase was attended by the lecturers who answered questions and helped with the functionality of Pundit. Due to technical problems the comparative exploration of the created statements was only possible as a principle demonstration by the lecturers.

The prepared annotation vocabulary constitutes an attempt to translate a methodological approach to the historical and critical analysis of historical images, in this case digitised photographs, to a simple and flat ontology. The method translated was based on several methodological approaches and the expertise of the teacher. In the annotation vocabulary we differentiated the following levels in the analysis of images: (1) the context of provenance including information about the author and the

historical context of creation, the shown things in the image (Bezugsrealität), the used stylistics in the image (Bildrealität), and the historical and personal perception of the image (Wirkungsrealität).

During the translation process several conceptual issues arose including the following. The social aspects which need to be considered are potentially unlimited and deciding on relevance on particular aspects in advance is not feasible. The solution was to use generic properties and classes which allow to either create your own textual information (Literals) or by creating your own specific instances, i.e. create your own terminological system. The same is true in the case of existing interpretation offerings and one's own semantics, where we also resorted to generic annotation entities. For example, we introduced the property "wirkt" (has effect) along with a class for personal and existing interpretative impressions. Here, students were able to create their own instance data with Korbo.

Even though the translation process posed more conceptual obstacles than previous ones, the result nethertheless proved to constitute an applicable and already useful attempt to represent an interpretative approach in the formal and explicit Linked Data structure. The teacher stated also one principal issue with regard to the digital working as a whole that the haptic aspect of the physical image would be lost during and also that spontaneous in interpretative process might not be adequately covered by computers and their formal and explicit working mode.

2. Questionnaire

This section summarises the results from the questionnaire for all three experiments. Only where appropriate the discussion will differentiate between the individual use cases. Sections which are specific to the individual use cases have been reported in the previous sections.

The questionnaire was taken on the last day of each workshop and had 31 respondents.

2.1 Digital Humanities and Linked Data

The participants were first asked whether they encountered the terms Digital Humanities and Linked (Open) Data before the experiments. 26% (8) of the participants had heard of the term “Digital Humanities” before the experiment while 23 (74%) did not. Only 16% (5) knew of the term “Linked (Open) Data” before the experiment, and 84% (26) did not. Not surprisingly, 48% (15) would not call themselves “Digital Humanist”, and 39% (12) were not sure, while only 13% (4) would say that they are “Digital Humanists”. Accordingly, none of the participants has used a tool for semantic annotations before.

The participants were then asked which advantages they see in Linked (Open) Data tools for their own work. Most participants mentioned the facilitation of information integration, reusability and accessibility of information, i.e. of research data such as the annotations and the research objects and their relation to other objects. Another important advantage seen by the participants as a result of their work with Pundit and Linked Data was the more intensive and different engagement with the research object itself, and that the annotation vocabulary helped to work in a structured and systematic way. The possibility to explore the annotations in Ask was helpful to discover new relations and, at the same time, helped to keep an overview of the annotations created. More generally, some participants pointed out that the terminological system used during the experiment, i.e. the annotation vocabulary and the instance data, creates background knowledge which can be exchanged and reused in working groups.

The disadvantages of Linked (Open) Data tools for their own work were mostly related to the specific issues resulting from the current state of Pundit. The issue mentioned

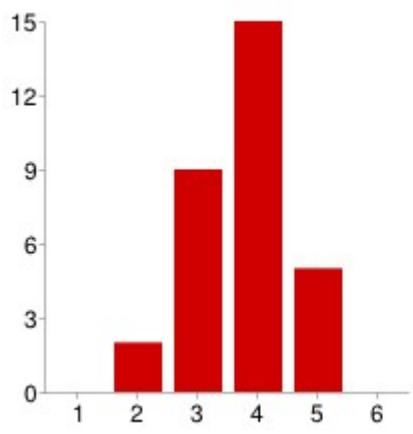
most often was the amount of time it takes to create triples, that too many triples result in complexity which is difficult to filter and possible redundancy of statements, and the necessity to communicate with developers to implement new classes and properties. Some participants mentioned the necessity to learn and understand the principles of Linked Data as another potential hindrance.

Participants from the FHP experiment also pointed out that no workflows exist yet, leading again to time consuming work processes, especially that there is no established means and workflows for quality checking before making something publicly available.

Next, the participants were asked in more detail about the functionality provided by Pundit and Linked Data in relation to research in the humanities.

2.2 Functionality of Pundit

The participants of the experiments found their experience with Pundit mostly positive, as shown in figure 1. 64% (20) rated their experience as rather positive while 25% (11) rated their experience as rather negative. Considering the current state of Pundit and its components which are not yet optimised for efficient and fluent workflows, this is surprisingly positive.



1	**0**	0%
2	**2**	6%
3	**9**	29%
4	**15**	48%
5	**5**	16%
6	**0**	0%

Figure 1: How would you rate your experience with Pundit? (1 = "very bad", 6 = "excellent")

The next question asked for the circumstances under which participants would use Pundit for their own research or work. Mostly, participants demanded general **better usability** during triple creation, better and intuitive interface in Pundit, and an overall more stable system.[11] These issues related to the current state of Pundit and its components were complemented by requests for additional features such as the option to apply an annotation to several pages at once, for example date or creator to the single pages of a letter, more space to enter free text (Literals), entity extraction, and in particular more and better filter options for annotations. Furthermore, options for access restrictions, for example to Korbo in case of instance data about persons were mentioned as important, and easy export of annotation to other software for further processing. Some participants mentioned support by professionals for creating vocabularies and training with this kind of software and working mode.

The last question in this section inquired about the stages in the research and work progress which participants would like to see supported by Pundit. The participants mostly indicated that one of the main application scenarios relates to researching and collecting facts on one's own research objects (personal collection), either collaboratively in groups or for individual research. In particular, **initial stages of the research process** where research objects are pre-analysed and the researcher tries to establish an overview on the corpus, such as formal or outer analysis of sources, appear to be a concrete and immediate application scenarios. The immediate value is seen in structuring and systematizing knowledge and to create structured research corpus with connected research objects which also allow to quickly retrieve sources by searching for entities such as persons, places, topics etc. Few participants explicitly mentioned the possibility to use the results from such initial phases for testing a

[11] Servers were down during the HUB experiment for a short time duration, and response times of the annotations servers were lagging occasionally due to the relatively high network traffic caused by the workshops.

vocabulary for opinion mining applications, or more generally preparing bigger projects and analysis.

The next section focused on various aspects of the digital and non-digital publication behaviour of the participants.

2.3 Publication

Roughly half of the respondents indicated to work in rather analogue settings while the rest indicated to work in rather digital settings. Only one respondent said to work only analogue while none said to work only digital. Most respondents are grouped in the middle of the scala. These results indicate that genuine digital working settings or contexts are not yet considerably established in the normal working routines of students.

On the other hand, regarding the question whether the respondents would publish their work digitally, 65% (20) answered with yes and 35% (11) answered with no. This indicates that publishing digital is slightly more common than working digitally, i.e. the process of research leading to a publication of research results is less affected by a digital setting.

When asked as to what digital publications are, most respondents provided a broad range of general answers such as any kind of document or information accessible or available online such e-journals, qualification theses via edoc-repositories or websites, Web portals, either as open access or with some access restriction such as pay barrier. Few considered ebooks, music files, or videos files as digital publications as well, while only one named digitised objects available via Europeana or the Deutsche Digitale Bibliothek as digital publications.

The question regarding under which circumstances the participants would publish digitally generated very diverse aspects. All respondents appeared to have considered only traditional text-based publications such as articles here which coincides with the previous responses. Many respondents seem to prefer publication as open access or at least by retaining various rights such as whether one retains the copyright or the right to decide where else the publication will be made available. Few would publish completely freely without any restrictions, while more respondents would allow free access and use limited to academia or particular communities. Other important aspects mentioned were legal consideration regarding copyright and privacy laws in cases of documents related to individuals and the reputation of the publication channel or platform

The annotations created in Pundit were considered as a publication by roughly half of the respondents, 48% (15), while 29% (9) answered no, and 23% (7) were not sure. At the same time, 84% (26) would make their annotations created with Pundit public and available to others. Only 3% (1) would not, and 13% (4) were undecided.

The reasons provided for being willing to make annotations available were manifold. Most reasons were concerned with the potential usefulness for other users. For example, providing annotations on things such as persons or places to others would deliver additional context knowledge on research objects and might help others with search and retrieval. In this context, crowd-sourcing for collecting contextual information on research objects is seen as a kind of fruitful publication. Furthermore, sharing annotations is seen as having the potential to facilitate research and to compare results and to gain feedback in order to improve one’s own annotation data.

However, several concerns and reservations were expressed: Few argued that as long as annotations are not really (re-)useable, for example being citable and referenceable, they cannot be considered to be publications, or that annotations per se are no proper

publications and always need support by proper text, that annotations are only supportive to research and more of a collaborative endeavour.

Some would only want to publish “factual” annotations but no “subjective” annotations which are based on interpretative acts or which are personal comments or notes. One respondent made the distinction regarding the content of the annotation: if it is basic, simple information, then it is less important to understand an annotation as a proper publication but if it is more high-level content expressed by the annotation then it is very important. Problematic is also that the reasoning leading to the triple is not obvious and missing which could be problematic. If annotations are considered proper publications then a quality check of annotations would be necessary before publication since the correctness of the annotations is important.

Several respondents raised concerns regarding whether they would retain the rights on their annotations or if potential employer would hold the right on the annotations made as part of a working contract. Uncertainty existed regarding the violation of rights or copyrights of annotated research objects such as digitised photographs or archival material, or personal privacy. Another issue raised was the question what it means to reuse a single entity (resource) from an external knowledge base in a triple: Who would “own” the complete annotation then?

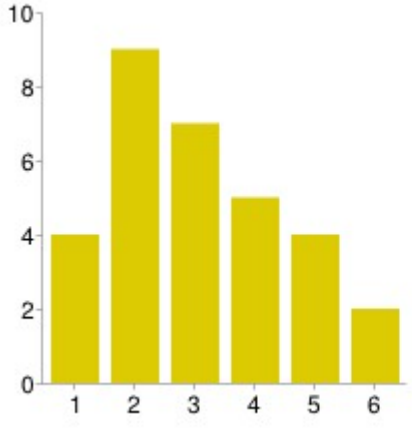
Participants seemed to associate publication of triples as making them accessible in some way but appear not to have pondered about more integrated and contextualised forms of publishing triple data, for example as part of a documented package of statements about a text. This could mean that participants do not consider triples as proper and publishable research. Publication appears to be interpreted as a means for collaboration and supporting each other in research. And so, even though a majority of the respondents would consider annotations as some kind of publication and most

respondents would make their annotations available to others, most participants of the experiments were very aware of the potential legal, social, and technical issues surrounding the publication of annotations.

The next section in the questionnaire inquired about annotations.

2.4 Annotations

The participants were asked whether they experienced the triple structure of the annotations as restrictive. The responses are spread but the respondent tend to find the triple structure of the annotations as unrestrictive, as shown in figure 2.



1	4	13%
2	9	29%
3	7	23%
4	5	16%
5	4	13%
6	2	6%

Figure 2: Do you find the structure of the annotations (triple) as restrictive? (1 = "I absolutely not agree", 6 = "I completely agree"

Some of the reasons given for experiencing the triple structure as restrictive include the missing ability for adapting the vocabulary during the course of the annotation work. New statements which are found as necessary during the subsequent work in Pundit cannot be created without extending the existing vocabulary. Pundit, in its current state, does not allow easy extension of the vocabulary because it requires editing and knowledge of configuration files in JSON. In this regard, this is a limitation of the current system. However, in terms of Linked Data principles, extending vocabularies at any

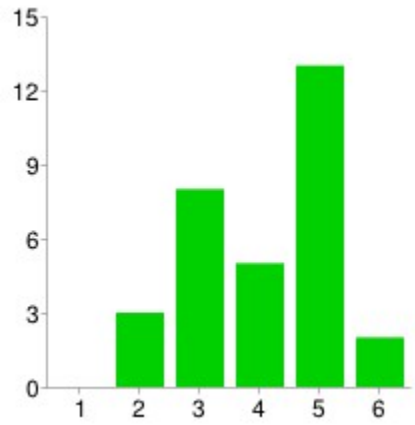
time, i.e. not systematically and with rigor and caution, may quickly lead to an influx of statements and schema entities.

Some respondents feared losing information because of missing classes or properties in the vocabulary, or because of the missing possibility to make differentiated statements such as weighted statements. Several respondents criticised that they have to create many triples in order to express slightly more complex information. These participants then also stated that they had to put these background information into free text fields but, at the same time, realised that these free text information are not accessible to further processing.

In general, there appears to be unease and uncertainty towards understanding which annotations are best created as triples and which information is better suited for free text fields (Literals).

However, after having overcome the hurdle of becoming familiar with the principal approach of Linked Data, a majority of the respondents recognised and appreciated the structured approach to annotation as “logical” and systematic. Some advantages stated were that a structured collection of research data evolves and that the annotation vocabulary provides guidance to what should be annotated.

While a slight majority of the respondents tend to feel not restricted by the triple structure of the annotations, a clearer majority considered the provided vocabularies as being adequate for the tasks they had to perform during the experiments (cf. “Use Cases”), as shown in figure 3.



1	**0**	0%
2	**3**	10%
3	**8**	26%
4	**5**	16%
5	**13**	42%
6	**2**	6%

Figure 3: In your opinion, were the provided vocabularies adequate? (1 = "not adequate at all", 6 = "completely adequate")

The next question asked how vocabularies should be created and managed: 23% (7) prefer to create and manage the vocabularies on their own, 39% (12) prefer to collaborate with other scientists from the same field, 19% (6) prefer to collaborate with the developers of the tools, and 10% (6) prefer to only re-use existing vocabularies and to leave the creation and management to the developers of the tools.

Various reasons were given in the follow-up question. Most respondents tend to favour a collaborative approach including either several other researchers or additionally developers. On the one hand, the single researchers know their domain and research objects best and therefore know which kinds of statements or extensions they would need. Developers alone would not be able to foresee all relevant classes and properties. On the other hand, however, the danger of losing semantic interoperability of one's own research data is also seen if the researchers would be allowed to freely manage or successively extend their vocabularies. Therefore, many respondents stressed the importance to collaborate with developers which would help to retain rigor in the vocabulary but also with other scientists in order to avoid too specialised vocabularies or to identify missing entities.

Other responses from FHP also suggested that in larger institutions or working groups, such as archives or divisional departments, selected people could coordinate and manage the vocabulary based on the feedback from researchers in order to retain rigor and also to quickly adapt to new projects. On the other hand, one respondent feared that such collaborative approaches could be too time consuming if larger groups would have to agree on modifications in vocabularies and then depend on implementation by a third party. Being able to add new entities would facilitate the working process.

All in all, the respondents tend to favor collaborative approaches involving researchers and developers to the creation and management of shared vocabularies and stress the importance to be able to specialise their vocabularies in such a context.

The last question inquired for potential reuse scenarios of the personal triples created by the participants. 35% (11) of the respondents could not think of potential reuse scenarios for their triple data. The other respondents mostly indicated three different kinds of principal reuse scenarios: The reuse of the annotations by other researchers working either with the same or similar research objects or on similar research questions. Reuse of previous annotations would be time saving. Some respondents explicitly pointed out that the reuse of annotations, i.e. research data, would be equal to considering previous research. The second principal reuse scenario mentioned by the respondents was using annotations as additional contextual information for search and retrieval allowing, for example, access via person concepts to images. Few respondents pointed visualisation out as a third scenario for further processing the annotations in other programs.

2.5 Ontology

In the questionnaire, each use case had a dedicated section with questions tailored towards the particular use case. These questions focused on the application of the

specific annotation vocabularies and feedback regarding the principal usefulness of the Linked Data approach for the particular domain covered by the uses case.

UBER

The participants from **UBER** mostly stated that they were able to make the most relevant annotations. However, one of the major problems were annotations which would demand to express uncertainty about particular statements, for example, saying that a photograph has been probably taken at a particular time of day. Similarly, expressing assumptions or reflexion, or creating something like a footnote, was not possible but often demanded. Generally, some participants wished for more properties to describe more of the historical background, i.e. context information which is not directly related to the annotated resource itself.

Problematic statements were about the authenticity of photographs, whether a statement about the semantic structure of a photograph relates to the whole image or only a fragment or both.

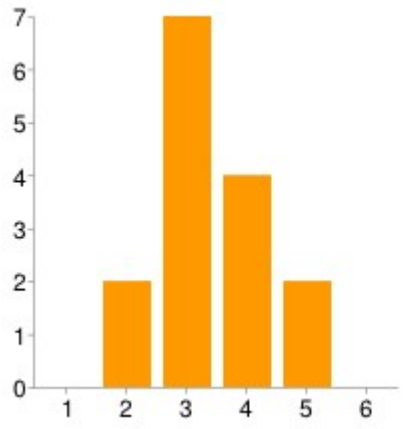
In general, the more interpretation was necessary during the analysis of a photograph (semantic and symbolic structure mostly) the more difficult it became for the participants to reduce these interpretations to factual statements and concepts. Since there was no possibility to express assumptions or uncertainty, many participants chose to not create respective annotations. Furthermore, some participants found the structure of the triple itself as a source for uncertainty because natural language statements lose all grammar.

Some participants stated that the reflection on the process of analysis of a photograph was sharpened by the forced explicitness of the vocabulary and triple structure. In this context, their creativity and inspiration had been aided by the instance data of others available through Korbo but also in Linked Open Data sources.

Apart from comments on the current functionality of Pundit and Ask, which is too time consuming and cumbersome, several remarks were made regarding potential disadvantages for digital critique of images. Some fear to loose information due to the restrictive triple structure. The reason most likely is due to the inability to easily add new properties (and classes) to the vocabularies during the work. Similarly, another mentioned issue is that the interpretative acts are lost in the triple structure, the reason why an annotation has been made, and that, potentially, the creativity of these interpretative acts is lost. Lastly, the annotations do not have any scientific reliability in so far as there are no established measurements for such a purpose.

Some participants had the impression that too much information is being created during the annotation so that the overview quickly diminished. This impression is certainly due to the limited facilities of Pundit and Ask to easily filter annotations. Being able to freely create triples might also entail the danger to lose focus on the actual objective of the current working task because you can go on with triples “in any direction”. Other feared that too many allowed and publicly available perspectives and opinions - expressed through annotations and instances created by users - could lead to a lot of wrong or bad information.

The final question in this section inquired whether the UBER participants think they overall successfully worked with Pundit and the annotation vocabulary on the research questions and tasks. As figure 4 shows, respondents tended to judge the success more sceptical than the participants from FHP. Considering the complex and difficult topic operationalised for this use case, this is not bad.



1	**0**	0%
2	**2**	6%
3	**7**	23%
4	**4**	13%
5	**2**	6%
6	**0**	0%

Figure 4: Would you say that you successfully worked on the research questions and work tasks with Pundit and the provided ontology?

FHP

In the case of the FHP experiment, most participants were able to create all relevant statements. Examples of missing statements include provenance relations, i.e. relation between document and holding or collection, and the relation to the archive, structural relations such as next page, and various details such as nicknames of persons, gender, additional information about places, or a dedicated property for describing the content of a text. Nearly all of these shortcomings were solved by the participants by using the free comment property. All specified missing statements would be easy to provide from existing ontologies. In the context of the experiment, time constraints prohibited adding appropriate entities to the vocabulary.

Most participants indicated that they were able to make all statements they deemed necessary during the workshops. One participant specifically asked for the possibility to create class hierarchies in order to allow clearer ontological differentiation between class and instance. Few participants felt that the German labels for the properties in particular were not always fitting in terms of the articles. Some asked for better class and property descriptions.

Some participants missed an option to create complex statements, for example by combining several triples. The facility of Pundit to create templates for RDF statements was not used in the experiments but might have been a possibility to address this issue. Few specific, conceptual issues were raised such as how to deal with the changing borders and the location of historical places. The target of annotations in the context of specific statements was not always clear to the participants. For example, should place names or person names be annotated by using the complete document or the particular text-fragment as the subject of the triple. Another general question was where the line should be drawn between the information the archivist and the information the researcher should annotate.

As advantages the often mentioned aspects were stated such as collaborative work, additional context for documents, the re-usability of annotation data, and connecting and integrating information.

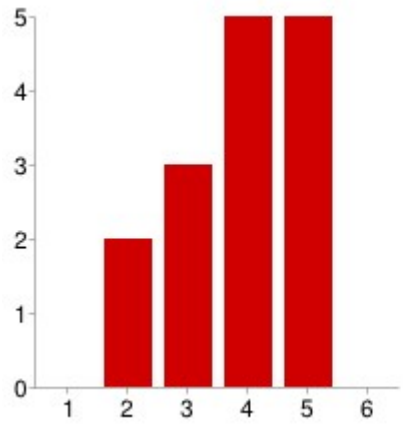
Potential disadvantages were that, at least in the context of Pundit, every annotation can be edited or deleted leading to potentially unstable and unreliable research data, i.e. the annotations. Furthermore, there is no reliable system for checking and sustaining the integrity and authenticity of statements. A related question is how trust in terms of the scientific reliability of the Linked Data sources could be established.

Even though collaboration has often been mentioned as positive, i.e. in terms of sharing the workload or re-using information, some participants also pointed out that collaboration may also lead to confusion and potentially dissent during work. Also, if collaboration is understood as involving people from outside a group or organisation, this could mean losing control over the documentation process if anyone would be allowed to annotate.

Some respondents stressed the importance of having stable and accepted vocabularies in order to have relevant annotations and avoiding too much information. Achieving this aim, however, demands a lot of work and coordination before productive annotation could be done.

Also, the necessary training regarding Linked Data principles and tools and the possible additional work in terms of coordination and quality control have been mentioned as potential obstacles.

FHP participants were also asked how they would judge the success they had in translating and applying the editorial guideline to Linked Data and Pundit. The responses are fairly positive, as shown in figure 5, which confirms the general tendency of the responses provided in the other questions.



1	**0**	0%
2	**2**	6%
3	**3**	10%
4	**5**	16%
5	**5**	16%
6	**0**	0%

Figure 5: Would you say that the translation and application of the editorial guideline into the Linked Data format was successful?

FHP respondents were further asked about potential usage scenarios for Linked Data in the context of digital editions and archives. Most respondent said that Linked Data could be used, in principle, for guidelines for digital editions but also stressed that proper

vocabularies, documentation, and workflows are necessary, and that the actual tools need to be easier to use.

In the case of archives, most respondents were skeptical regarding the application of Linked Data and annotations with tools such as Pundit in archival contexts and especially the daily work of archives. The Linked Data approach in combination with annotations may have potential for (explorative) search functionality in archives allowing to discover sources connected to particular entities, or, more general, to provide additional contextual information on particular documents or holdings.

The biggest obstacle, however, is the lack of resources in archives to learn and employ such techniques and tools, the mass of documents, and the issue of access restricted holdings. Regarding the latter, participants called for proper technical and administrative policies which would ensure access restriction to triples describing such holdings. Therefore, in the short-term, the Linked Data approach may be mostly useful in the context of small and closed projects, for example for presenting and publishing collections of documents about a topic, but not for daily archival work. In the future, the Linked Data approach could potentially be used for the description stage in the archival workflow where information such as call numbers or content description are added to archival material.

Lastly, in the case of editing source documents, most respondents were also relatively sceptical. A principal advantage mentioned is the flexible terminological system consisting of instances created within a structured and stable framework, the vocabulary. The workshop showed that the annotation already worked well for small comments and editing small remarks, but is less useful for longer transcriptions. In general, some respondents felt that the annotations are too detached from the edited document.

GEI

The GEI participant first provided feedback on the annotation vocabulary and the annotation process.

During the work process, additional properties seemed to be necessary. During the interpretative work with the school books several layers of the interpretation became apparent. Statements differ in their level of specificity and how directly they relate to phenomena which are immanent to the text. These different layers should be incorporated into the annotation vocabulary in order to differentiate them. Most problematic were statements where the participant was uncertain about the statement: Differentiating how sure he was about a statement was not possible but deemed very important.

However, the participant was not sure whether new properties should be introduced due to potential conceptual and technical issues such as too many properties which could diminish re-usability. The respondent stressed that sufficient time for preparation is very important to prepare a stable and conclusive vocabulary. However, all in all, the vocabulary proved to useful and sufficient for most relevant statements.

The participant would refrain from creating annotation vocabularies alone but would prefer to do so in collaboration with developers. The Linked Data and Semantic Web principles, for example, is not easy to understand and the participant would prefer to have explanation for important terms provided by experts.

Furthermore, the possibility to create hierarchies of classes and properties is important in order to create a proper terminological system for annotation. Here, the current functionality of Pundit prohibited to create class taxonomies for the annotation vocabulary, however, RDFS does provide the necessary means for building such taxonomies.

Lastly, the participant would have preferred to have more options for a more flexible visualisation of the created statements. The respondent preferred a chronological visualisation of the annotations. Even though the faceted browser allowed to explore the statements created, a possibility to have a more intuitive chronological representation of the statements would have been preferred. For example, showing all triples including a connotation chronological and in the context of the schoolbook the annotation was made in.

Regarding the potential usefulness of Pundit and Linked Data approaches, the participant identified as potential useful application scenarios of Linked Data and Pundit collaboration within research groups and projects where work task could be easily assigned to different people and the results then merged afterwards. For example, assistants could be given the annotation vocabulary as a guideline and catalogue of criteria by which they would search and describe documents in an archive such as historical journals or newspapers. The annotations created by using the vocabulary would allow assessing, comparing, and merging the results from the different assistants more easily.

In this context, another advantage was seen in the fact that the annotations vocabulary and the created instance data, the annotations, are independent from the application. However, Pundit is missing an easy option to export this data in order to reuse it in different contexts. For example, other applications could reuse the create statements and annotation vocabulary as a basis for automated analysis of a large corpus of schoolbooks.

Major problems and disadvantages of Pundit and Linked Data were, according to the participant, the relative high effort necessary to create annotations. This process is too

time consuming at the moment in terms of the usability of the annotation tool Pundit, where it takes too many clicks to create an annotation.

There was unease with the technical infrastructure on which the participant felt dependent in so far as he would not be able to work anymore if the server or the Internet connection broke down. The participant would prefer some kind of backup or intermediate or local temporary storage which would allow to work at any time and independently from a working internet connection or server. Changes would be synced back to the server when being online again.

Another potential problem pointed out are legal issues. The data used for this use case is under control of the GEI. However, in general, the participant stated that it is unclear to him how exactly copyright and reuse of digital sources, digitised documents or Linked Data instances, are handled and what that means for one's own research work.

The GEI participant judged the general success of the work conducted in the context of Pundit and Linked Data as more positive (scale of 4).

3. Conclusion

The experiments were conducted within the limitations imposed by the current state of Pundit which primarily include the inability to create class taxonomies in annotation vocabularies and the limitation to faceted browsing in tabular formats based on the entities incorporated in the vocabulary and instance data (schema and instance elements). Despite these limitations, the results of the experiments underline the usefulness of the principal approach of Pundit and of Linked Data for genuine research questions and interests of humanists by providing evidence for their applicability in the context of interpretative approaches in the humanities.

In each use case, we were able to create **useable drafts** of annotation vocabularies in RDF(S) for the respective research interests in relatively short time: for the analysis of connotations in historical schoolbooks, the historical critique and analysis of photographs, and for editorial work on archival documents. Each annotation vocabulary remained simple and without complex ontological constructs. Even though several interpretative processes were difficult to explicate and even more difficult to formalize, the iterative translation process is already a genuine part of the interpretative research process. This iterative process of translating methodological approaches and interpretative statements, not only during preparatory stages, but continuously during the actual annotation phase, needs to be considered much more intensively and systematically than was possible in the context of these experiments. New and necessary statements will appear only during the application of the annotation vocabulary and need to be fed back into the vocabulary. For constructs which were not representable in the annotation vocabularies, informal conventions have been introduced, for example, by prescribing particular textual values for objects in triples in cases where information was either unknown or uncertain.

However, all annotation vocabularies proved to be reasonably **productive for their respective purposes**. In this regard, simple annotation vocabulary in RDF(S) appear to be able to support very **different kinds of research interests**, in the context of the experiment, that is the archival-historical domain.

In particular, simple Linked Data annotation vocabularies proved to be relatively **accessible** to humanists. Even though all participants had no prior working knowledge of Linked Data or Pundit, they were mostly able to grasp the concept of triple annotations within a couple of hours and thereafter apply the annotations vocabularies. In this regard, results appear to be obtainable for students with relatively low

prerequisites which is an important aspect for lowering the access barrier to Linked Data annotation tools.

Furthermore, the formal and explicit approach of Linked Data appears to have initiated reflection on participants' own working practices. Student were forced to reflect on their own work processes because of the explicitness of the vocabulary. The annotation vocabularies and the created instance data provide a common basis for discussion on the method, what should be said, and interpretation, what has been said, of the research objects in particular contexts since Korbo and faceted browsers allow to explore any triple data relatively easily. On the other hand, teachers have a pedagogical tool for communicating theoretical and practical representations of methods. In so far, the annotation vocabularies and Pundit constitute a potential epistemological tool for educational contexts.

The reuse of the created annotation vocabularies in other similar use cases remains open since much work would have to be invested into their further specification in order to be useable by other researchers. However, the experiments could also be seen as a first round of evaluation and testing of such vocabularies providing the basis for future refinements. In order to be able to develop "real" application scenarios a translation between the requirements of the researchers and their respective research process, on the one hand, and the functionalities of the virtual research environment developed for them, on the other hand, must happen in an iterative translation process.

Future experiments with Pundit or similar tools would need to prepare more specific adaptations to the single use cases, especially in terms of flexible visualisation and filter options which proved to be the most pivotal incentive and means for establishing a sense of usefulness for proper and sustained research. Chronological visualisations and immediate contexts are important and comparison of entities.

The experiments resulted in several recommendations regarding the general research question of Task 3.4 “How can Linked Data based digital tools and data support, facilitate, or enhance humanist work practices?”. The following list contains the most relevant recommendations:

1. Providing clear **mission statements** towards the purpose of a digital tool and vocabulary in relation to the overall research process and particular stages is essential. For example, which stage or segment of the research process does an annotation vocabulary address? The general point has been made in the context of the SDM already: One of the most important aspects is communication between the humanist and the developers.
2. In this regard, the aspect of good **usability** of tools and interface needs to be stressed again. Without ease of use and efficiency, tools such as Pundit will not be used voluntarily in any serious or productive scenario. Tools need to provide feedback at any point why a particular functionality does not work.
3. Tools and workflows need to be established, in order to implement and maintain a collaborative and iterative **development and application of the annotation vocabularies** for interpretative approaches. This includes a vocabulary browser and editor which allows to edit and extend vocabularies by classes, properties and scope notes in a controlled and flexible manner. How appropriate policies and workflows should look like remains subject to future research.
4. For interpretative applications of annotations vocabularies it is important to investigate how to **qualify statements**, for example, allowing to provide reasons or to express uncertainty, and to ensure proper provenance information. Interpretative statements are the result of proper research work and need to be attributable and citable. Furthermore, **access rights management** for

statements is important where access to statements can be regulated granularily. These are difficult topics which need to addressed, however, in order to start establishing scientific reliability and trust.

5. The **stability** of the annotated research objects and instance data is another crucial aspect. If the annotated objects or instance data used in triples from Linked Open Data knowledge bases disappear, the result of the interpretative research work is rendered invalid. This issue has been raised by participants. Furthermore, the teachers pointed out that they have to grade the students also based on the triple statements they created. Here, sustainability of these triples is crucial and as is a clear provenance of statements.
6. In this context, the option to **export** the research data, i.e. the created annotations and instance data from the notebooks to locally stored files is important. Even though not all participants raised the issue where annotations are currently stored, or were even conscious of the problem, when directly asked about the issue, the need to export data was stressed. Motivations were to reuse the data in other applications such as for more sophisticated visualisations, to have full control of the data, and to have a backup of the data available. This will contribute to more trust in digital tools regarding the safety of the personal research data.
7. Lastly, **legal issues** of using and reusing Web resources of any kind including Linked Data resources, need to be discussed, explained and communicated. Even though users of the Web are aware that the same or similar legal constraints and formal and informal obligations pertain to the use of digital resources as in the analogue world, uncertainty about the exact ramifications, rules, and regulations prevail. Several examples have been given in the previous discussion pertaining to the use of images but also to Linked Data resources such as the legal status and intellectual rights towards annotations and even

single resources, whether thumbnails are already duplications of the original picture, or the *Schöpfungshöhe* of annotations. The legal status of these matter to humanists.

8. Approach the students and young researchers outside the “Digital Humanities”!

For future research and further development of workflows for interpretative approaches implemented with semantic annotations tools such as Pundit, we suggest the following 3-tiered process. Even though partially predetermined by Pundit, the three steps proved to be useful and an appropriate approach to engage humanists with Linked Data in a way which, provided appropriate implementation of the necessary tools, will gradually allow to improve the usefulness for humanist research. The same basic iterative 3-tiered process appeared to be valid in the context of the reasoning experiment (cf. “Reasoning”) where interpretative modelling has been investigated in the context of Pundit and Linked Data.

We therefore propose an iterative 3-tiered process to be implemented and offered to humanists in order to enable them to begin with meaningful and useful interpretative work with Linked Data enable semantic annotation environments:

1. *Conceptualising*: Selecting, modifying or creating a vocabulary (referential structure) for annotation is already a genuine part of the research process. This process needs to involve both sides, the humanists and developers. A method and policy to revisit and modify the initial vocabulary needs to be implemented since necessary statements develop during research work, the annotation work.
2. *Annotating*: Applying the annotation vocabulary to a research object is the second step. During this phase, humanists appear to prefer to create their own referential data, either because necessary instance data does not exist or is not being trusted (the Linked Data knowledge bases appear alien to the participants). If they import instance data they prefer to have a filter on import in order to have

control over what is being imported into their notebook. Important is also to allow statements regarding the rationale and weighting of a statement which was one of the major demanded features.

3. *Visualising*: The process of exploring what has been created in terms of the annotation vocabulary, statements and instance data. Here, humanists want to apply their own “reasoning” by filtering and adopting the visualization context to their needs. Relevant generic visualisation were the simple comparison of two or more entities of the same type and their immediate attributes, and chronological ordering and displaying of statements.

Acknowledgments

The following members and associates of the Digital Humanists Advisory Board (DHAB) contributed to the development of the Scholarly Domain Model and the writing of this paper: Markus Schnöpf (BBAW/FHP), Andreas Weiß (GEI), Sabine Moller (UBER)